REVSTAT – Statistical Journal
Volume 14, Number 3, June 2016, 245–271

A MIXTURE INTEGER-VALUED GARCH MODEL

Authors: Mamadou Lamine Diop
– LERSTAD, Université Gaston Berger, Saint Louis, Sénégal
diopml@yahoo.r

Aliou Diop
– LERSTAD, Université Gaston Berger, Saint Louis, Sénégal
aliou.diop@ugb.edu.sn

Abdou Kâ Diongue
– LERSTAD, Université Gaston Berger, Saint Louis, Sénégal
abdou.diongue@ugb.edu.sn

Received: July 2013 Revised: July 2014 Accepted: November 2014

Abstract:

- In this paper, we generalize the mixture integer-valued ARCH model (MINARCH) introduced by Zhu *et al.* (2010) (F. Zhu, Q. Li, D. Wang. A mixture integer-valued ARCH model, *J. Statist. Plann. Inference*, 140 (2010), 2025–2036.) to a mixture integer-valued GARCH (MINGARCH) for modeling time series of counts. This model includes the ability to take into account the moving average (MA) components of the series. We give the necessary and sufficient conditions for first and second order stationarity solutions. The estimation is done via the EM algorithm. The model selection problem is studied by using three information criterions. We also study the performance of the method via simulations and include a real data application.

Key-Words:

- *integer-valued; mixture models; GARCH; EM algorithm.*

AMS Subject Classification:

- 62F10, 62M10.

1. INTRODUCTION

Time series count data are widely observed in real-world applications (epidemiology, econometrics, insurance). Many different approaches have been proposed to model time series count data, which are able to describe different types of marginal distribution. Zeger (1988) discusses a model for regression analysis with a time series of counts by illustrating the technique with an analysis of trends in U.S. polio incidence, while Ferland *et al.* (2006) proposed an integer-valued autoregressive conditional heteroscedastic (INARCH) model to deal with integer-valued time series with overdispersion. Zhu (2011) proposed a negative binomial INGARCH (NBINGARCH) model that can deal with both overdispersion and potential extreme observations simultaneously. Zhu (2012) introduced a generalized Poisson INGARCH model, which can account for both overdispersion and underdispersion, among others.

In the literature, time series are often assumed to be driven by a unimodal innovation series. However, many time series may exhibit multimodality either in the marginal or the conditional distribution. For example, Martin (1992) proposed to model multimodal jump phenomena by a multipredictor autoregressive time series (MATS) model, while Wong and Li (2000) generalized the GMTD model to the full mixture autoregressive (MAR) model whose predictive distribution could also be multimodal. Muller and Sawitzki (1991) proposed and studied a method for analyzing the modality of a distribution.

Recently, Zhu *et al.* (2010) have used the idea of Saikkonen (2007) on the definition of a very general mixture model to generalize the INARCH model to the mixture (MINARCH) model, which has the advantages over the INARCH model because of its ability to handle multimodality and non-stationary components. But, they did not take into account the MA part of the model. Sometimes, as in the classical GARCH model, large number of lagged residuals must be included to specify the model correctly. As it is well known that computational problem may arise when the autoregressive polynomial in the conditional mean of the MINARCH model presents high order, we introduce in this paper the MINGARCH model which is a natural generalization of the MINARCH model.

The paper is organized as follows. In Section 2 we describe the MINGARCH model and the stationarity conditions. In Section 3, we discuss the estimation procedures by using an expectation-maximization (EM) algorithm introduced by Dempster *et al.* (1997) with a simulation study. We illustrate the usefulness of the model in Section 4 by an empirical example. A brief discussion and concluding remarks are given in Section 5.

2. THE MIXTURE INTEGER-VALUED GARCH MODEL

The MINGARCH(K; $p_1, ..., p_K$; $q_1, ..., q_K$) model is defined by:

$$\left\{ \begin{array}{l} X_t = \sum_{k=1}^{K} \mathbb{1}(\eta_t = k) Y_{kt}\,, \\ Y_{kt}|\mathcal{F}_{t-1} : \mathcal{P}(\lambda_{kt})\,, \\ \lambda_{kt} = \alpha_{k0} + \sum_{i=1}^{p_k} \alpha_{ki} X_{t-i} + \sum_{j=1}^{q_k} \beta_{kj}\,\lambda_{k(t-j)}\,, \end{array} \right. \tag{2.1}$$

where $\mathcal{P}(\lambda)$ is the Poisson distribution with parameter λ, $\alpha_{k0} > 0$, $\alpha_{ki} \geq 0$, $\beta_{kj} \geq 0$, $(i = 1, ..., p_k,\ \ j = 1, ..., q_k,\ \ k = 1, ..., K)$, $\mathbb{1}(\cdot)$ denotes the indicator function, p_k and q_k are respectively the MA and AR orders of λ_{kt}, $\mathcal{F}_{t-1}$ indicates the information given up to time $t-1$, η_t is a sequence of independent and identically distributed random variables with $\mathbb{P}(\eta_t = k) = \alpha_k$, $k = 1, ..., K$. It is assumed that X_{t-j} and η_t are independent for all t and $j > 0$, the variables Y_{kt} and η_t are conditionally independent given $\mathcal{F}_{t-1}$, $\alpha_1 \geq \alpha_2 \geq ... \geq \alpha_K$ for identifiability (see Titterington (1985)) and $\sum_{k=1}^{K} \alpha_k = 1$. If $\beta_{kj} = 0$, $k = 1, ..., K$, $j = 1, ..., q_k$, the model is denoted MINARCH(K; $p_1, ..., p_K$).

The MINGARCH model is able to handle the conditional overdispersion in integer-valued time series. In fact, the conditional mean and variance are given by

$$\mathbb{E}\Big(X_t|\mathcal{F}_{t-1}\Big) = \sum_{k=1}^{K} \alpha_k\,\lambda_{kt}\,,$$

and

$$\mathrm{Var}\Big(X_t|\mathcal{F}_{t-1}\Big) = \mathbb{E}\Big(X_t|\mathcal{F}_{t-1}\Big) + \sum_{k=1}^{K} \alpha_k\,\lambda_{kt}^2 - \left(\sum_{k=1}^{K} \alpha_k\,\lambda_{kt}\right)^2.$$

Using the Jensen's inequality, we can easily see that:

$$\sum_{k=1}^{K} \alpha_k\,\lambda_{kt}^2 - \left(\sum_{k=1}^{K} \alpha_k\,\lambda_{kt}\right)^2 > 0\,.$$

Hence the conditional variance is greater than the conditional mean. Furthermore

$$\begin{aligned} \mathrm{Var}\Big(X_t\Big) &= \mathbb{E}\Big(\mathrm{Var}\big(X_t|\mathcal{F}_{t-1}\big)\Big) + \mathrm{Var}\Big(\mathbb{E}\big(X_t|\mathcal{F}_{t-1}\big)\Big) \\ &= \mathbb{E}\left(\sum_{k=1}^{K} \alpha_k\,\lambda_{kt} + \sum_{k=1}^{K} \alpha_k\,\lambda_{kt}^2 - \left(\sum_{k=1}^{K} \alpha_k\,\lambda_{kt}\right)^2\right) + \mathrm{Var}\left(\sum_{k=1}^{K} \alpha_k\,\lambda_{kt}\right) \\ &= \mathbb{E}\Big(X_t\Big) + \sum_{k=1}^{K} \alpha_k\,\mathbb{E}\Big(\lambda_{kt}^2\Big) - \Big(\mathbb{E}\big(X_t\big)\Big)^2 \\ &\geq \mathbb{E}\Big(X_t\Big) + \mathbb{E}\left(\sum_{k=1}^{K} \alpha_k\,\lambda_{kt}^2 - \left(\sum_{k=1}^{K} \alpha_k\,\lambda_{kt}\right)^2\right). \end{aligned}$$

Then the variance is larger than the mean, which indicates that model (2.1) is also able to describe the time series count with overdispersion.

Let us now introduce the polynomials $D_k(B) = 1 - \beta_{k1}B - \beta_{k2}B^2 - \cdots - \beta_{kq}B^q$, $k = 1, ..., K$, where B is the backshift operator. In the following, we assume that

H_1 : For $k = 1, ..., K$, the roots of $D_k(z) = 0$ lie outside the unit circle,

H_2 : $\lambda_{kt} < \infty$ a.s. for any fixed t and k.

Let $p = \max(p_1, ..., p_K)$; $q = \max(q_1, ..., q_K)$; $\alpha_{ki} = 0$, for $i > p_k$; $\beta_{kj} = 0$, for $j > q_k$ and $L = \max(p, q)$.

First and second-order stationarity conditions for the MINGARCH model (2.1) are given in Theorem 2.1 and Theorem 2.2.

Theorem 2.1. *Assume that the conditions H_1 and H_2 hold. A necessary and sufficient condition for model (2.1) to be stationarity in the mean is that the roots of the equation:*

$$1 - \sum_{k=1}^{K} \alpha_k \left(\frac{\sum_{i=1}^{p_k} \alpha_{ki} Z^{-i}}{1 - \sum_{j=1}^{q_k} \beta_{kj} Z^{-j}} \right) = 0 \tag{2.2}$$

lie inside the unit circle.

Proof: Let $\mu_t = \mathbb{E}(X_t) = \sum_{k=1}^{K} \alpha_k \mathbb{E}(\lambda_{kt})$ for all $t \in \mathbb{Z}$. Since

$$\lambda_{kt} = \alpha_{k0} + \sum_{i=1}^{p_k} \alpha_{ki} X_{t-i} + \sum_{j=1}^{q_k} \beta_{kj} \lambda_{k(t-j)},$$

the recursion equation gives, for all $m > 1$,

$$\begin{aligned} \lambda_{kt} = {} & \alpha_{k_0} + \sum_{i=1}^{L} \alpha_{ki} X_{t-i} + \sum_{l=1}^{m} \sum_{j_1,...,j_l=1}^{L} \alpha_{k0} \beta_{kj_1} \cdots \beta_{kj_l} \\ & + \sum_{l=1}^{m} \sum_{j_1,...,j_{l+1}=1}^{L} \alpha_{kj_{l+1}} \beta_{kj_1} \cdots \beta_{kj_l} X_{t-j_1-\cdots-j_l-j_{l+1}} \\ & + \sum_{j_1,...,j_{m+1}=1}^{L} \beta_{kj_1} \cdots \beta_{kj_{m+1}} \lambda_{k(t-j_1-\cdots-j_{m+1})}. \end{aligned}$$

Let $C_{k0} = \alpha_{k0} + \sum_{l=1}^{\infty} \sum_{j_1,...,j_l=1}^{L} \alpha_{k0} \beta_{kj_1} \cdots \beta_{kj_l}$. We define

$$\lambda'_{kt} = C_{k0} + \sum_{i=1}^{L} \alpha_{ki} X_{t-i} + \sum_{l=1}^{\infty} \sum_{j_1,...,j_{l+1}=1}^{L} \alpha_{kj_{l+1}} \beta_{kj_1} \cdots \beta_{kj_l} X_{t-j_1-j_2-\cdots-j_{l+1}}. \tag{2.3}$$

Since $\sum_{j=1}^{L}\beta_{kj}<1$ it is easy to see that $0\le\lambda'_{kt}<\infty$ a.s. for any fixed t and k.

We will show below that $\lambda_{kt}=\lambda'_{kt}$ almost surely as $m\to\infty$ for any fixed t and k. In what follows, C will denote any positive constant whose value is unimportant and may vary from line to line. Let t and k be fixed now. It follows that for any $m\ge 1$

$$\begin{aligned}\left|\lambda_{kt}-\lambda'_{kt}\right| \le & \sum_{l=m+1}^{\infty}\sum_{j_1,\dots,j_l=1}^{L}\alpha_{k0}\,\beta_{kj_1}\cdots\beta_{kj_l}\\ &+\sum_{l=m+1}^{\infty}\sum_{j_1,\dots,j_{l+1}=1}^{L}\alpha_{kj_{l+1}}\beta_{kj_1}\cdots\beta_{kj_l}\,X_{t-j_1-j_2-\cdots-j_{l+1}}\\ &+\sum_{j_1,\dots,j_{m+1}=1}^{L}\beta_{kj_1}\cdots\beta_{kj_{m+1}}\,\lambda_{k(t-j_1-\cdots-j_{m+1})}\,.\end{aligned}$$

Under H_2, we have

$$E\left\{\sum_{j_1,\dots,j_{l+1}=1}^{L}\alpha_{kj_{l+1}}\beta_{kj_1}\cdots\beta_{kj_l}\,X_{t-j_1-j_2-\cdots-j_{l+1}}\right\}\le C\left(\sum_{j=1}^{L}\beta_{kj}\right)^{l},$$

and

$$E\left\{\sum_{j_1,\dots,j_{m+1}=1}^{L}\beta_{kj_1}\cdots\beta_{kj_{m+1}}\,\lambda_{k(t-j_1-\cdots-j_{m+1})}\right\}\le C\left(\sum_{j=1}^{L}\beta_{kj}\right)^{m+1}.$$

The expectation of the right-hand side of the above is bounded by

$$\left(C_{k0}+C_1\left(1-\sum_{j=1}^{L}\beta_{kj}\right)^{-1}\right)\left(\sum_{j=1}^{L}\beta_{kj}\right)^{m+1}.$$

Let $A_m=\left\{|\lambda_{kt}-\lambda'_{kt}|>\frac{1}{m}\right\}$. Then

$$\mathbb{P}(A_m)\le m\left(C_{k0}+C_1\left(1-\sum_{j=1}^{L}\beta_{kj}\right)^{-1}\right)\left(\sum_{j=1}^{L}\beta_{kj}\right)^{m+1}.$$

Then, using Borel–Cantelli lemma and the fact that $A_m\subset A_{m+1}$, we can show that $\lambda_{kt}=\lambda'_{kt}$ a.s. Therefore,

$$\begin{aligned}(2.4)\qquad \mu_t =& \sum_{k=1}^{K}\alpha_k\,C_{k0}+\sum_{i=1}^{L}\sum_{k=1}^{K}\alpha_k\,\alpha_{ki}\,\mu_{t-i}\\ &+\sum_{l=1}^{\infty}\sum_{j_1,\dots,j_{l+1}=1}^{L}\sum_{k=1}^{K}\alpha_k\,\alpha_{kj_{l+1}}\beta_{kj_1}\cdots\beta_{kj_l}\,\mu_{t-j_1-j_2-\cdots-j_{l+1}}\,.\end{aligned}$$

The necessary and sufficient condition for a non-homogeneous difference equation (2.4) to have a stable solution, which is finite and independent of t, is that all roots of the equation

$$1-\sum_{i=1}^{L}\sum_{k=1}^{K}\alpha_k\,\alpha_{ki}\,Z^{-i}-\sum_{l=1}^{\infty}\sum_{j_1,\ldots,j_{l+1}=1}^{L}\sum_{k=1}^{K}\alpha_k\,\alpha_{kj_{l+1}}\,\beta_{kj_1}\cdots\beta_{kj_l}\,Z^{-(j_1+j_2+\cdots+j_{l+1})}=0$$

lie inside the unit circle (see Goldberg (1958)). This equation is equivalent to

$$1-\sum_{k=1}^{K}\alpha_k\left(\sum_{i=1}^{p_k}\alpha_{ki}\,Z^{-i}\right)\sum_{l=0}^{\infty}\left(\sum_{j=1}^{q_k}\beta_{kj}\,Z^{-j}\right)^{l}=0\,.$$

Since $\sum_{j=1}^{q_k}\beta_{kj}<1,\ k=1,...,K$ and $\|Z\|<1$, the equation (2.2) follows. □

As an illustration, we consider in the following corollary the MINARCH(K; $p_1,...,p_K$).

Corollary 2.1. *A necessary and sufficient condition for the MINARCH(K; $p_1,...,p_K$) model to be first-order stationary is that the roots of the equation*

$$1-\sum_{i=1}^{p}\left(\sum_{k=1}^{K}\alpha_k\,\alpha_{ki}\right)Z^{-i}=0$$

lie inside the unit circle, where $p=\max(p_1,...,p_K)$.

Now, we consider the MINGARCH model with $p_k=q_k=1$ for all $k=1,...,K$. The following corollary gives a necessary and sufficient condition for the MINGARCH($K;1,...,1;1,...,1$) model to be stationary in the mean.

Corollary 2.2. *A necessary and sufficient condition for the MINGARCH($K;1,...,1;1,...,1$) model to be first-order stationarity is that the roots of the equation*

$$1+C_1Z^{-1}+C_2Z^{-2}+\cdots+C_KZ^{-K}=0$$

lie inside the unit circle where

$$C_1=-\sum_{k=1}^{K}(\delta_k+\alpha_k\gamma_k)$$

and

$$C_j=(-1)^j\left[\sum_{k_1>k_2>...>k_j}^{K}\delta_{k_1}\delta_{k_2}\cdots\delta_{k_j}+\sum_{k=1}^{K}\alpha_k\gamma_k\left(\sum_{\substack{k_1>k_2>...>k_{j-1}\\k_1\neq k,\,k_2\neq k,\,...,\,k_{j-1}\neq k}}^{K}\delta_{k_1}\delta_{k_2}...\delta_{k_{j-1}}\right)\right]$$

for $j=2,...,K$, with $\gamma_k=\alpha_{k1}$ and $\delta_k=\beta_{k1}$.

Proof: The equation (2.2) becomes

$$1-\sum_{k=1}^{K}\frac{\alpha_k\gamma_k Z^{-1}}{1-\delta_k Z^{-1}} = 0\,.$$

Reducing to the same denominator, the preceding equation is equivalent to:

$$\prod_{k=1}^{K}\left(1-\delta_k Z^{-1}\right)-\sum_{k=1}^{K}\alpha_k\gamma_k Z^{-1}\prod_{\substack{k'=1\\k'\neq k}}^{K}\left(1-\delta_{k'} Z^{-1}\right) =$$
$$= 1+C_1Z^{-1}+C_2Z^{-2}+\cdots+C_KZ^{-K} = 0\,. \qquad \square$$

From equation (2.4), we have

$$\mathbb{E}(X_t) = \mu = \sum_{k=1}^{K}\alpha_k C_{k0}+\mu\sum_{l=0}^{\infty}\sum_{j_1,\ldots,j_{l+1}=1}^{L}\sum_{k=1}^{K}\alpha_k\,\alpha_{kj_{l+1}}\beta_{kj_1}\cdots\beta_{kj_l}\,. \tag{2.5}$$

Hence

$$\mu = \frac{\sum_{k=1}^{K}\left(\frac{\alpha_k\,\alpha_{k0}}{1-\sum_{j=1}^{q_k}\beta_{kj}}\right)}{1-\sum_{k=1}^{K}\left(\frac{\sum_{i=1}^{p_k}\alpha_k\,\alpha_{ki}}{1-\sum_{j=1}^{q_k}\beta_{kj}}\right)}\,.$$

A necessary condition for first-order stationarity of model (2.1) is given in the following proposition.

Proposition 2.1. *Under conditions H_1 and H_2, a necessary condition for first-order stationarity of model (2.1) is*

$$\sum_{k=1}^{K}\left(\frac{\sum_{i=1}^{p_k}\alpha_k\,\alpha_{ki}}{1-\sum_{j=1}^{q_k}\beta_{kj}}\right) < 1\,.$$

Remark 2.1.

1. As a special case, a necessary condition for the MINGARCH$(2;1,1;1,1)$ model to be stationary in the mean is:

$$\frac{\alpha_1\alpha_{11}}{1-\beta_{11}}+\frac{\alpha_2\,\alpha_{21}}{1-\beta_{21}} < 1\,.$$

2. When the process (X_t) follows a MINARCH$(K;p_1,\ldots,p_K)$, the condition stated in Proposition 2.1, reduced $\sum_{k=1}^{K}\left(\sum_{i=1}^{p_k}\alpha_k\,\alpha_{ki}\right)<1$ as in Zhu *et al.* (2010).

The second order stationarity condition for the MINGARCH model (2.1) in given the following theorem. Its proof is postponed in an Appendix.

Theorem 2.2. *Let $(X_t)_{t\in\mathbb{Z}}$ be a MINGARCH$(K; p_1, ..., p_K; q_1, ..., q_K)$ model. Assume that the conditions H_1 and H_2 hold. If the process $(X_t)_{t\in\mathbb{Z}}$ is first-order stationary then a necessary and sufficient condition for the process to be second-order stationary is that all roots of $1 - c_1 Z^{-1} - c_2 Z^{-2} - \cdots - c_L Z^{-L} = 0$ lie inside the unit circle, where*

$$c_u = \sum_{k=1}^{K} \alpha_k \left(\Delta_{k,u} - \sum_{v=1}^{L-1} \Lambda_{kv}\, b_{vu}\, \omega_{u0} \right), \quad u = 1, ..., L-1 \text{ and } c_L = \sum_{k=1}^{K} \alpha_k\, \Delta_{k,L}\,,$$

with

$$\Delta_{k,i} = \Delta_{k,i}^{(1)} + \Delta_{k,i}^{(2)}\,,$$

$$\Delta_{k,i}^{(1)} = \sum_{l=0}^{\infty} \sum_{\substack{j_{l+2}=i \\ j_{l+2}=j_1+\cdots+j_{l+1}}}^{L} \alpha_{kj_{l+1}} \alpha_{kj_{l+2}}\, \beta_{kj_1} \cdots \beta_{kj_l}\,,$$

$$\Delta_{k,i}^{(2)} = \sum_{\substack{l=0 \\ l'=0}}^{\infty} \sum_{\substack{j_1+\cdots+j_{l+2}=i \\ j_1+\cdots+j_{l+2}=j'_1+\cdots+j'_{l'+1}}}^{L} \alpha_{kj_{l+2}}\, \beta_{kj_1} \cdots \beta_{kj_{l+1}}\, \alpha_{kj'_{l'+1}} \beta_{kj'_1} \cdots \beta_{kj'_{l'}}\,,$$

$$\Lambda_{kv} = \Lambda_{kv}^{(1)} + \Lambda_{kv}^{(2)}\,,$$

$$\Lambda_{kv}^{(1)} = \sum_{l=0}^{\infty} \sum_{|j_{l+2}-j_1-\cdots-j_{l+1}|=v}^{L} \alpha_{kj_{l+1}} \alpha_{kj_{l+2}}\, \beta_{kj_1} \cdots \beta_{kj_l}\,,$$

$$\Lambda_{kv}^{(2)} = \sum_{\substack{l=0 \\ l'=0}}^{\infty} \sum_{\left|j_1+\cdots+j_{l+2}-j'_1-\cdots-j'_{l'+1}\right|=v}^{L} \alpha_{kj_{l+2}}\, \beta_{kj_1} \cdots \beta_{kj_{l+1}}\, \alpha_{kj'_{l'+1}} \beta_{kj'_1} \cdots \beta_{kj'_{l'}}\,,$$

and $\Gamma = (\omega_{ij})_{i,j=1}^{L-1}$, $\Gamma^{-1} = (b_{ij})_{i,j=1}^{L-1}$, two matrices such that

$$\omega_{i0} = \sum_{l=0}^{\infty} \sum_{k=1}^{K} \alpha_k\, \delta_{i0kl}\,, \quad \omega_{iu} = \sum_{l=0}^{\infty} \sum_{k=1}^{K} \alpha_k\, \delta_{iukl} \text{ for } u \neq i\,, \quad \omega_{ii} = \sum_{l=0}^{\infty} \sum_{k=1}^{K} \alpha_k\, \delta_{iikl} - 1\,,$$

$$\delta_{iukl} = \sum_{|i-j_1-\cdots-j_{l+1}|=u} \alpha_{kj_{l+1}}\, \beta_{kj_1} \cdots \beta_{kj_l}\,.$$

We remark that when (X_t) follows a MINARCH$(K; p_1, ..., p_K)$, Theorem 2.2 reduces to Theorem 2 of Zhu *et al.* (2010), where $L = \max(p_1, ..., p_K)$.

If the process (X_t) following a MINGARCH$(K; p_1, ..., p_K; q_1, ..., q_K)$ model is second-order stationary, then from (5.2), we have

$$\mathbb{E}(X_t^2) = \frac{c_0}{1 - \sum_{u=1}^{L} c_u}\,,$$

where $c_0 > 0$ (see Appendix B). Hence, necessary second order stationary condition for a special case is given by the following proposition.

Proposition 2.2. *The second order stationary condition for a MINGARCH*$(K; 1, ..., 1; 1, ..., 1)$ *is* $c_1 < 1$ *where* $c_1 = \sum_{k=1}^{K} \alpha_k \, \alpha_{k1}^2$.

In the following theorem, we give a necessary and sufficient condition for the process (X_t) following a MINGARCH$(K; 1, ..., 1; 1, ..., 1)$ model to be m order stationary. The results for the general model MINGARCH$(K; p_1, ..., p_K; q_1, ..., q_K)$ are difficult to obtain and need further investigations.

Theorem 2.3. *The* m*-th moment of a MINGARCH*$(K; 1, ..., 1; 1, ..., 1)$ *model is finite if and only if*

$$\sum_{k=1}^{K} \alpha_k \left(\alpha_{k1} + \beta_{k1}\right)^m < 1 \,. \tag{2.6}$$

Proof: Since $Y_{kt}|\mathcal{F}_{t-1}$ is a Poisson random variable with mean $\lambda_{kt} = \alpha_{k0} + \alpha_{k1} X_{t-1} + \beta_{k1} \lambda_{k(t-1)}$ conditionally to time $t-1$, the m-th moment of X_t is given by

$$\mathbb{E}\left(X_t^m\right) \;=\; \sum_{k=1}^{K} \alpha_k \sum_{j=0}^{m} \left\{ {m \atop j} \right\} \mathbb{E}\big(\lambda_{kt}^j\big)$$

where $\left\{ {m \atop j} \right\}$ is the Stirling number of the second kind (see Gradshteyn and Ryzhik (2007), p. 1046) and

$$\lambda_{kt}^j \;=\; \sum_{n=0}^{j} \binom{j}{n} \alpha_{k0}^{j-n} \sum_{r=0}^{n} \binom{n}{r} \left(\alpha_{k1} X_{t-1}\right)^r \left(\beta_{k1} \lambda_{k(t-1)}\right)^{n-r} .$$

We mimic the proof of Proposition 6 in Ferland *et al.* (2006) by setting

$$\mathbf{\Lambda}_{k,t} \;=\; \left(\lambda_{kt}^m, ..., \lambda_{kt}^2, \lambda_{kt}\right)^T$$

and showing that for all k

$$\mathbb{E}(\mathbf{\Lambda}_{k,t}|\mathcal{F}_{t-2}) \;=\; \mathbf{d}_k + \mathbf{D}_k \mathbf{\Lambda}_{k,t-1}$$

where $\mathbf{d}_k$ and $\mathbf{D}_k$ are respectively a constant vector and an upper triangular matrix. The derivation of the required condition follows the great lines of the proof of Proposition 6 in Ferland *et al.* (2006). □

The result obtained in Theorem (2.6) is an extension of Proposition 6 in Ferland *et al.* (2006) for an INGARCH$(1, 1)$ process. When the β_{ki}'s equal to zero, the (necessary) condition (2.6) is a special case of the result obtained in Theorem 3 of Zhu *et al.* (2010).

3. PARAMETER ESTIMATION AND SIMULATION

3.1. Estimation procedure

In this section, we discuss the estimation of the parameters of a MINGARCH model by using the expectation-maximization (EM) algorithm (see Dempster *et al.* (1997)). Suppose that the observation $X = (X_1, ..., X_n)$ is generated from the MINGARCH model.

Let $Z = (Z_1, ..., Z_n)$ be the random variable where $Z_t = (Z_{1,t}, ..., Z_{K,t})^T$ is a vector whose components are defined by:

$$Z_{i,t} = \begin{cases} 1 & \text{if } \; X_t \text{ comes from the } i\text{-th component; } 1 \leq i \leq K\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

The vectors Z_t are not observed and its distribution is:

$$\mathbb{P}\big(Z_t = (1, 0, ..., 0)^T\big) = \alpha_1\,, \quad ..., \quad \mathbb{P}\big(Z_t = (0, 0, ..., 0, 1)^T\big) = \alpha_K\,.$$

Let $\alpha = (\alpha_1, ..., \alpha_{K-1})^T$, $\alpha_{(k)} = (\alpha_{k0}, \alpha_{k1}, ..., \alpha_{kp_k})^T$, $\beta_{(k)} = (\beta_{k1}, ..., \beta_{kq_k})^T$ $\theta_{(k)} = (\alpha_{(k)}^T, \beta_{(k)}^T)$ and $\theta = (\alpha, \theta_{(1)}, ..., \theta_{(K)})^T \in \Theta$ (the parameter space).

Given Z_t, the distribution of the complete data (X_t, Z_t) is then given by

$$\prod_{k=1}^{K} \left(\alpha_k \frac{\lambda_{kt}^{X_t} \exp(-\lambda_{kt})}{X_t!}\right)^{Z_{kt}}.$$

Let l_t be the conditional log-likelihood function at time t. The log-likelihood is given by $l(\theta) = \sum_{t=1}^{n} l_t$.

$l(\theta)$ is the joint log-likelihood function of the first L random variables of the series and $l^*(\theta) = \sum_{t=L+1}^{n} l_t$ is called the conditional log-likelihood function. When the sample size n is large, the influence of $\sum_{t=1}^{L} l_t$ will be negligible. In this study, the parameters will be estimated by maximizing the conditional log-likelihood function l^* given by

$$l^*(\theta) = \sum_{t=L+1}^{n} \left\{ \sum_{k=1}^{K} Z_{kt} \log(\alpha_k) + X_t \sum_{k=1}^{K} Z_{kt} \log(\lambda_{kt}) - \sum_{k=1}^{K} Z_{kt}\lambda_{kt} - \log(X_t!) \right\}.$$

The first derivatives of the conditional log-likelihood with respect to θ are:

$$\text{(3.1)} \qquad \frac{\partial l*}{\partial \alpha_k} = \sum_{t=L+1}^{n} \left(\frac{Z_{kt}}{\alpha_k} - \frac{Z_{Kt}}{\alpha_K}\right), \qquad k = 1, ..., K-1\,,$$

$$(3.2)\qquad \frac{\partial l*}{\partial \alpha_{ki}} = \sum_{t=L+1}^{n} Z_{kt}\frac{X_t-\lambda_{kt}}{\lambda_{kt}}\,U(X_t,i)\,, \qquad k=1,...,K,\quad i=0,...,p_k\,,$$

$$(3.3)\qquad \frac{\partial l^*}{\partial \beta_{kj}} = \sum_{t=L+1}^{n} Z_{kt}\frac{X_t-\lambda_{kt}}{\lambda_{kt}}\,\lambda_{k,t-j}\,, \qquad k=1,...,K,\quad j=1,...,q_k\,,$$

where $U(X_t,i)=1$ for $i=0$ and $U(X_t,i)=X_{t-i}$ for $i>0$.

Given that the process $\{Z_t\}$ is not observed, the data that we have do not allow the estimation of the parameter θ. An iterative procedure (EM) is proposed for estimating the parameters by maximizing the conditional log-likelihood function $l^*(\theta)$, which consists of two steps (E and M) that we describe in the following.

E-step:

Suppose that θ is known. The missing data $\mathbf{Z}$ are then replaced by their conditional expectations, conditional on the parameters and on the observed data X. In this case the conditional expectation of the k-th component of Z_t is just the conditional probability that the observation X_t comes from the k-th component of the mixture distribution conditional on θ and X. Let $\tau_{k,t}$ be the conditional expectation of Z_{kt}.
Then the E-step equation is given by:

$$\tau_{k,t} = \frac{\alpha_k {\lambda_{kt}}^{X_t}\exp(-\lambda_{kt})}{\sum_{i=1}^{K}\alpha_i {\lambda_{it}}^{X_t}\exp(-\lambda_{it})}$$

where $k=1,2,...,K$ and $t=L+1,...,n$. In practice, we take $Z_{kt}=\tau_{k,t}$ from the previous E-step of the EM procedure.

M-step:

The missing data Z are replaced by their conditional expectations on the parameters θ and on the observed data $X_1,...,X_n$. The estimates of the parameters θ can then be obtained by maximizing the conditional log-likelihood function $l^*(\theta)$ by equating expressions (3.2)–(3.3) to 0.
The M-step equations become

$$\hat{\alpha}_k = \frac{1}{n-L}\sum_{t=L+1}^{n}\tau_{k,t}\,, \qquad k=1,...,K\,.$$

From the equation (3.2), we have:

$$\sum_{t=L+1}^{n}\frac{\tau_{t,k}X_t}{\hat{\lambda}_{kt}}\,U(X_t,i) = \sum_{t=L+1}^{n}\tau_{k,t}\,U(X_t,i)\,.$$

Then

$$\sum_{t=L+1}^{n}\left\{\frac{\tau_{k,t}\,X_t}{\sum_{j=0}^{p_k}\hat{\alpha}_{kj}\,U(X_t,j)+\sum_{j=1}^{q_k}\hat{\beta}_{kj}\hat{\lambda}_{k(t-j)}}\,U(X_t,i)\right\}=\sum_{t=L+1}^{n}\tau_{k,t}\,U(X_t,i)\,,$$

for $k=1,...,K,\quad i=0,...,p_k$.
Similarly equation (3.3) gives:

$$\sum_{t=L+1}^{n}\frac{\tau_{k,t}\,X_t}{\hat{\lambda}_{kt}}\,\hat{\lambda}_{k,t-j}\;=\sum_{t=L+1}^{n}\tau_{k,t}^{(s)}\,\hat{\lambda}_{k,t-j}^{(s)}\,.$$

Then

$$\sum_{t=L+1}^{n}\left\{\frac{\tau_{k,t}^{(s)}X_t}{\sum_{i=0}^{p_k}\hat{\alpha}_{ki}^{(s)}\,U(X_t,i)+\sum_{t=L+1}^{q_k}\hat{\beta}_{ki}^{(s)}\,\hat{\lambda}_{k,t-i}^{(s)}}\,\hat{\lambda}_{k,t-j}^{(s)}\right\}=\sum_{t=L+1}^{n}\tau_{k,t}^{(s)}\,\hat{\lambda}_{k,t-j}^{(s)}\,,$$

for $k=1,...,K,\;j=1,...,q_k$.

The estimate of θ is then obtained by iterating these two steps until convergence. The criterion used for checking convergence of the EM procedure is

$$\max\left\{\left|\frac{\theta_i^{(s+1)}-\theta_i^{(s)}}{\theta_i^{(s)}}\right|,\;s,i\geq 1\right\}\;\leq\;10^{-5}$$

where $\theta_i^{(s)}$ is the i-th component of θ obtained in the s-th iteration.

Among different strategies for choosing starting initial values for the EM algorithm (see Karlis and Xekalaki (2003), Melnykov and Melnykov (2012)), the random initialization method is employed in this paper (the initial values for $\theta_{(k)}$ are chosen randomly from a uniform distribution and the mixing proportions are generated from a Dirichlet distribution). The asymptotic properties are not treated in this paper but they have been studied by many authors. For example, Nityasuddhia and Böhning (2003) have studied the asymptotic properties of the EM algorithm estimate for normal mixture models. They show that the EM algorithm gives reasonable solutions of the score equations in an asymptotic unbiased sense. The performance of the EM algorithm is assessed by some simulation experiments.

3.2. Simulation studies

Monte Carlo experiment was conducted to investigate the performance of the EM estimation method. In all these simulation experiments, we use 100 independent realizations of the MINGARCH model defined in (2.1) with sizes $n = 100$, $n = 200$ and $n = 500$. The following two models were used in the experiment. The first, denoted Model (I), is a MINGARCH(2; 1,1; 1,1) model with parameter values

$$\begin{pmatrix} \alpha_1 & \alpha_{10} & \alpha_{11} & \beta_{11} \\ \alpha_2 & \alpha_{20} & \alpha_{21} & \beta_{21} \end{pmatrix} = \begin{pmatrix} 0.75 & 1.00 & 0.20 & 0.30 \\ 0.25 & 5.00 & 0.50 & 0.30 \end{pmatrix}.$$

The second, denoted Model (II), is a MINGARCH(3; 1,1,1; 1,1,1) model with parameter values

$$\begin{pmatrix} \alpha_1 & \alpha_{10} & \alpha_{11} & \beta_{11} \\ \alpha_2 & \alpha_{20} & \alpha_{21} & \beta_{21} \\ \alpha_3 & \alpha_{30} & \alpha_{31} & \beta_{31} \end{pmatrix} = \begin{pmatrix} 0.55 & 0.80 & 0.40 & 0.30 \\ 0.25 & 1.00 & 0.50 & 0.25 \\ 0.20 & 0.50 & 0.60 & 0.20 \end{pmatrix}.$$

The performances of the estimators are evaluated by the root mean square error (RMSE) and the mean absolute error (MAE).

Based on the results in Tables 1 and 2, we can see that as the sample size increases, the estimates seem to converge to the true parameter values.

Table 1: Results of the simulation study with model (I).

Sample size	k		α_k	α_{k0}	α_{k1}	β_{k1}
100	1	True values	0.7500	1.0000	0.2000	0.3000
		Mean estimated	0.7410	1.1883	0.1833	0.2446
		RMSE	0.0523	0.5789	0.0623	0.2137
		MAE	0.0405	0.4726	0.0506	0.1801
	2	True values	0.2500	5.0000	0.5000	0.3000
		Mean estimated	0.2590	5.1660	0.4619	0.2901
		RMSE	0.0523	2.6410	0.2823	0.2588
		MAE	0.0405	2.2060	0.2103	0.2274
200	1	True values	0.7500	1.0000	0.2000	0.3000
		Mean estimated	0.7463	1.0093	0.1909	0.3054
		RMSE	0.0359	0.4429	0.0468	0.1773
		MAE	0.0291	0.3641	0.0381	0.1460
	2	True values	0.2500	5.0000	0.5000	0.3000
		Mean estimated	0.2537	5.2571	0.4612	0.2928
		RMSE	0.0359	2.2616	0.1728	0.2380
		MAE	0.0291	1.8728	0.1314	0.1976
500	1	True values	0.7500	1.0000	0.2000	0.3000
		Mean estimated	0.7510	1.0646	0.1959	0.2817
		RMSE	0.0259	0.2525	0.0272	0.1035
		MAE	0.0212	0.1867	0.0214	0.0783
	2	True values	0.2500	5.0000	0.5000	0.3000
		Mean estimated	0.2490	5.3064	0.5026	0.2688
		RMSE	0.0259	1.6316	0.0982	0.1702
		MAE	0.0212	1.3483	0.0774	0.1443

Table 2: Results of the simulation study with model (II).

Sample size	k		α_k	α_{k0}	α_{k1}	β_{k1}
100	1	True values	0.5500	0.8000	0.4000	0.3000
		Mean estimated	0.5435	0.7671	0.4429	0.2163
		RMSE	0.1063	0.4997	0.1898	0.2339
		MAE	0.0828	0.4054	0.1482	0.1977
	2	True values	0.2500	1.0000	0.5000	0.2500
		Mean estimated	0.2240	1.0888	0.5344	0.2532
		RMSE	0.0802	0.7182	0.3804	0.2563
		MAE	0.0607	0.5504	0.2420	0.2113
	3	True values	0.2000	0.5000	0.6000	0.2000
		Mean estimated	0.2323	0.9516	0.4475	0.2714
		RMSE	0.0600	0.7127	0.2413	0.2263
		MAE	0.0429	0.5490	0.1895	0.1850
200	1	True values	0.5500	0.8000	0.4000	0.3000
		Mean estimated	0.5286	0.7471	0.4113	0.2552
		RMSE	0.1117	0.4363	0.1563	0.1942
		MAE	0.0838	0.3566	0.1190	0.1545
	2	True values	0.2500	1.0000	0.5000	0.2500
		Mean estimated	0.2316	1.0570	0.5340	0.2433
		RMSE	0.0785	0.6025	0.2584	0.1928
		MAE	0.0602	0.4787	0.1751	0.1506
	3	True values	0.2000	0.5000	0.6000	0.2000
		Mean estimated	0.2397	0.8867	0.4450	0.3042
		RMSE	0.0652	0.6088	0.2306	0.2439
		MAE	0.0452	0.4959	0.1806	0.1825
500	1	True values	0.5500	0.8000	0.4000	0.3000
		Mean estimated	0.5556	0.7040	0.4248	0.2725
		RMSE	0.0825	0.3246	0.1171	0.1797
		MAE	0.0614	0.2595	0.0934	0.1407
	2	True values	0.2500	1.0000	0.5000	0.2500
		Mean estimated	0.2182	0.9508	0.5223	0.2656
		RMSE	0.0620	0.4723	0.2059	0.2132
		MAE	0.0487	0.3853	0.1569	0.1576
	3	True values	0.2000	0.5000	0.6000	0.2000
		Mean estimated	0.2261	0.8985	0.4690	0.2815
		RMSE	0.0536	0.5883	0.1963	0.1988
		MAE	0.0298	0.4780	0.1586	0.1506

The performance of the estimate improves when the sample size increases. But this performance varies depending on the parameters. Indeed the parameter estimate α_k seems to give good results for all sample sizes considered. For the parameter α_{k0}, the RMSE and the MAE are slightly higher.

4. REAL DATA EXAMPLE

In this section we investigate the time series representing a count of the calls monthly reported in the 22nd police car beat in Pittsburg, starting in January 1990 and ending in December 2001. The data are available online at the forecasting principles site (`http://www.forecastingprinciples.com`), in the section about crime data. The summary statistics are given in Table 3. Mean and variance are estimated as 6.3056 and 23.0249, respectively. Hence the data seem to be overdispersed. The histogram of the series in Figure 1 shows that the series seems to be bimodal. Using the bimodality index of Der and Everitt (2002), Zhu *et al.* (2010) show that the series is bimodal. Moreover, they found that the MINARCH model is more appropriate for this dataset than the INARCH model. The autocorrelation function in Figure 2 implies that the moving average polynomial order is at most equal to three (i.e. $0 \leq q \leq 3$) while when considering the partial autocorrelation function, we can choose p such that $1 \leq p \leq 3$. Thus, in the following, we consider a MINGARCH model (2.1) with $K = 1, 2, 3$.

Table 3: Summary statistics of the crime counts series.

Sample size	Minimum	Maximum	Median	Mean	Variance	Skewness	Kurtosis
144	0	30	5	6.3056	23.0249	1.9732	8.7530

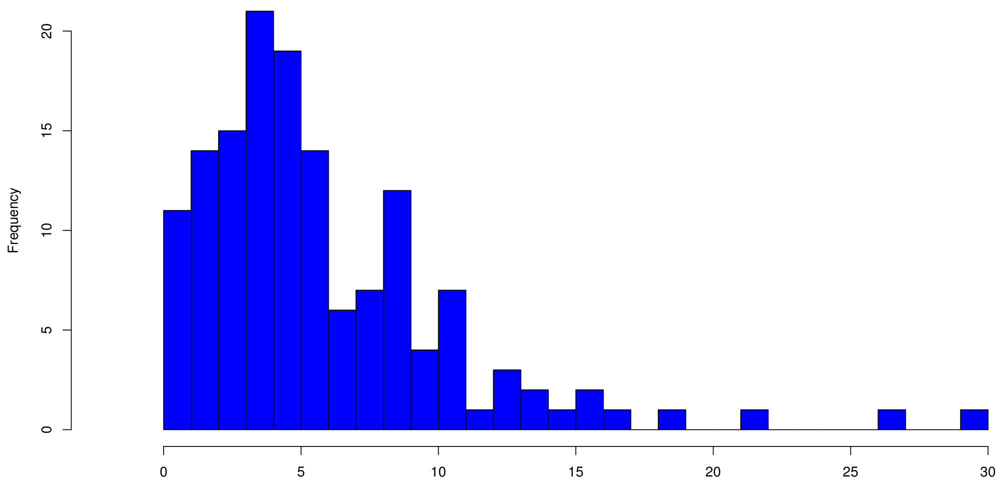

Figure 1: Histogram of the crime counts series.

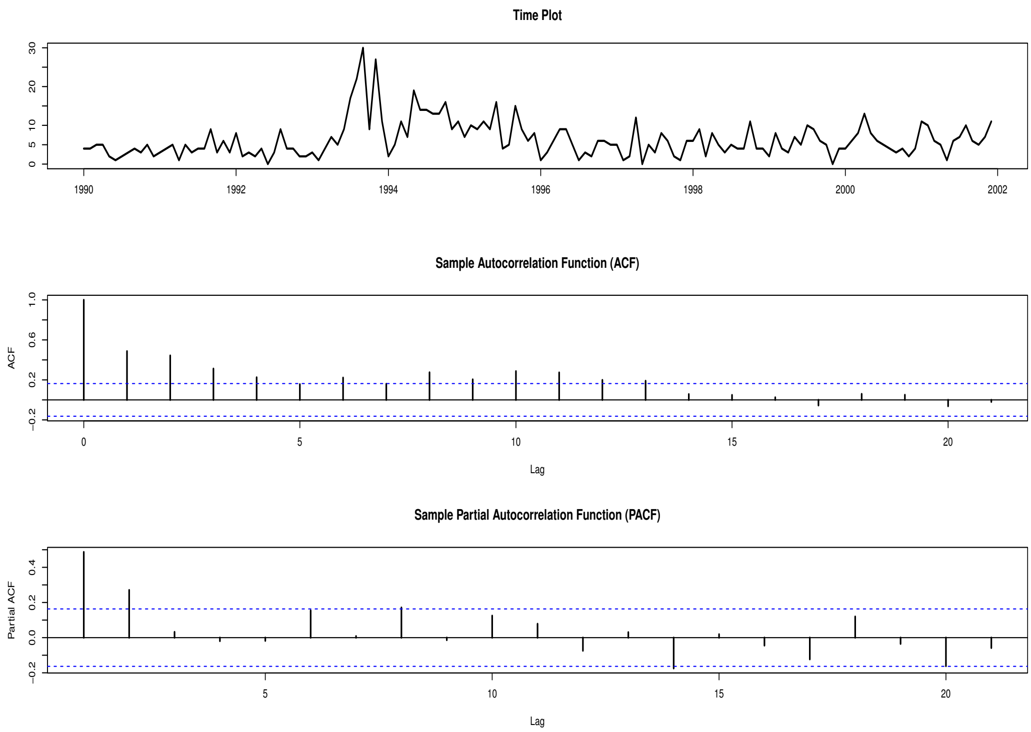

Figure 2: Crime counts series: the time plot, the sample autocorrelation and partial autocorrelation function.

The model selection criteria considered here are the Akaike information criterion (AIC), the Bayesian information criterion (BIC) and the mixture regression criterion (MRC) proposed by Naik *et al.* (2007). These two first criteria are both defined as minus twice the maximized log-likelihood plus a penalty term. The first choice is the maximum log-likelihood given by the EM estimation, it includes the information of the unobserved random variable $\mathbf{Z}$. The second choice is computed from the (conditional) probability density function of the MINGARCH model and is defined as

$$l' = \sum_{t=L+1}^{n} \log \left\{ \sum_{k=1}^{K} \alpha_k \frac{\lambda_{kt}^{X_t} \exp(-\lambda_{kt})}{X_t!} \right\} .$$

We use l' in this paper, it may have better performance in finite samples (see Wong and Li (2000)). The AIC and the BIC are given by:

$$\begin{aligned} AIC &= -2l' + 2\left(2K - 1 + \sum_{k=1}^{K} p_k + \sum_{k=1}^{K} q_k\right), \\ BIC &= -2l' + \log\Big(n - \max(p_{\max}, q_{\max})\Big)\left(2K - 1 + \sum_{k=1}^{K} p_k + \sum_{k=1}^{K} q_k\right). \end{aligned}$$

The MRC consists of three terms: the first measures the lack of fit, the second imposes a penalty for regression parameters, and the third is the clustering penalty function. For the MINGARCH model, the MRC is defined as

$$\mathrm{MRC}=\sum_{k=1}^{K}\widehat{n}_k\log(\widehat{\sigma}_k^2)+\sum_{k=1}^{K}\frac{\widehat{n}_k(\widehat{n}_k+\widehat{h}_k)}{\widehat{n}_k-\widehat{h}_k-2}-2\sum_{k=1}^{K}\widehat{n}_k\log(\widehat{\alpha}_k),$$

where $\widehat{n}_k=\mathrm{tr}(\widehat{W}_k)$, $\widehat{h}_k=\mathrm{tr}(\widehat{H}_k)$, $\widehat{\sigma}_k^2=(U-V\theta_k^*)^T\widehat{W}_k^{1/2}(I-\widehat{H}_k)(U-V\theta_k^*)/\widehat{n}_k$ with
$\widehat{W}_k=\mathrm{diag}\left((\widehat{\tau}_{k,L+1},...,\widehat{\tau}_{kn})^T\right)$, $\widehat{V}_k=\widehat{W}_k^{1/2}V$, $\widehat{H}_k=\widehat{V}_k\left(\widehat{V}_k^T\widehat{V}_k\right)^{-1}\widehat{V}_k^T$, $k=1,...,K$,
$V=(V_{L+1},...,V_n)^T$, $V_j=(1,X_{j-1},...,X_{j-p},\lambda_{k_j(j-1)},...,\lambda_{k_j(j-q)})^T$,
$k_j \mid \tau_{k_j,j}=\max\{\tau_{1,j},...,\tau_{K,j}\}$, $j=L+1,...,n$,
$\theta_k^*=\left(\alpha_{(k)}{}^T,\mathbf{0}^T,\beta_{(k)}{}^T,\mathbf{0}^T\right)^T_{(p+q+1)\times 1}$, $U=(X_{L+1},...,X_n)^T$.

The problem of model selection for MINGARCH models requires two aspects. First, we must select the number of components K. Second, the model identification problem needs to be addressed (i.e. the **AR** polynomial order, p_k, and the **MA** polynomial order, q_k). In this paper we not discuss the selection problem for the number of components. We concentrate on the order selection of each component. The order of the components is chosen to be that minimizing the values of the three criterions. The results are given in Tables 4, 5 and 6.

Table 4: AIC, BIC and MRC values for the crime counts series, $K=1$.

Order	AIC			BIC			MRC		
	$p=1$	$p=2$	$p=3$	$p=1$	$p=2$	$p=3$	$p=1$	$p=2$	$p=3$
$q=0$	832.52	813.55	811.50	838.44	**822.41**	823.29	558.37	546.64	**545.87**
$q=1$	815.11	813.62	813.31	824.01	825.44	852.80	550.56	548.57	547.16
$q=2$	813.04	815.06	813.39	824.87	829.84	831.09	548.57	550.50	549.37
$q=3$	**807.35**	809.35	811.82	846.83	827.04	832.46	546.31	548.31	547.19

Table 5: AIC, BIC and MRC values for the crime counts series, $K=2$.

Order	AIC			BIC			MRC		
	$p=1$	$p=2$	$p=3$	$p=1$	$p=2$	$p=3$	$p=1$	$p=2$	$p=3$
$q=0$	767.11	760.23	757.05	781.93	**780.92**	783.59	607.93	606.19	596.61
$q=1$	760.55	758.91	755.24	781.29	785.52	787.68	557.14	548.20	550.90
$q=2$	756.33	760.51	757.54	782.94	793.02	795.88	**537.59**	540.95	548.81
$q=3$	**751.82**	755.78	759.26	838.69	794.12	803.49	538.54	541.92	546.76

Table 6: AIC, BIC and MRC values for the crime counts series, $K = 3$.

Order	AIC			BIC			MRC		
	$p=1$	$p=2$	$p=3$	$p=1$	$p=2$	$p=3$	$p=1$	$p=2$	$p=3$
$q=0$	766.75	760.89	759.21	**790.46**	793.41	800.49	647.45	660.88	720.25
$q=1$	559.52	759.75	758.96	792.11	801.13	809.09	575.82	577.42	631.02
$q=2$	756.75	766.67	756.16	798.13	816.91	824.13	**573.34**	642.36	649.52
$q=3$	**749.04**	757.72	763.83	799.17	804.11	945.47	573.38	573.62	938.79

For the AIC, the BIC and the MRC, the minimums are represented by the underlined values. Based on the results in these tables (4, 5 and 6), the BIC and the MRC retain the two-component mixture model respectively with $(p, q) = (2, 0)$ and $(p, q) = (1, 2)$, which confirm the bimodality observed in the histogram. In contrast, the AIC retains the three-component mixture model with $(p, q) = (1, 3)$, which confirms the phenomena often observed in a lot of applications, namely that the AIC overclusters and overfits the data (for instance, see Naik *et al.* (2007)). In practice, it is observed that the BIC criterion selects the model of dimension smaller than the AIC criterion, which is not surprising since the BIC penalizes more than the AIC (when $n > 7$). We notice also that the next smallest AIC, BIC and MRC values are obtained in the two-component model with respectively $(p, q) = (1, 3)$, $(p, q) = (1, 1)$ and $(p, q) = (1, 3)$. These results confirm the result of the histogram and lends substantial support to the two-component model with $p = 1$ and $q \neq 0$. The values of the AIC and MRC obtained in our model are better than those of the MINARCH model. The values of BIC suggest the MINARCH$(2; 2, 2)$ model, but the smallest value is near of the BIC value obtained with MINGARCH$(2; 1, 1; 2, 2)$ model (780.92 and 782.94). In addition, the AIC of the MINGARCH$(2; 1, 1; 2, 2)$ (selected by the MRC) is better than the one in the MINARCH$(2; 2, 2)$ model. Hence, our results indicate that the MINGARCH model should be preferred to the MINARCH for this dataset.

5. CONCLUDING REMARKS

In this paper, a new model which generalizes the MINARCH model is proposed. Conditions for stationarity of the model and estimation procedure based on EM algorithm are investigated. Moreover, we study the finite performance of the estimation method using Monte Carlo simulations. Finally, a real case study is proposed. In a forthcoming, we plan to study the ergodicity conditions of the model as well as the optimal choice of the parameter K. In addition, we plan to study necessary and sufficient conditions for the MINGARCH$(K; p_1, ..., p_K; q_1, ..., q_K)$ process to be m order stationary for $m > 2$.

APPENDIX A — Proof of Theorem 2.2

Let $\gamma_{it} = E(X_t X_{t-i})$ for $i = 0, 1, ..., L$,

$$\begin{aligned}
\gamma_{it} &= \sum_{k=1}^{K} \alpha_k \mathbb{E}(\lambda_{kt} X_{t-i}) \\
&= \sum_{k=1}^{K} \alpha_{k0}\alpha_k E(X_{t-i}) + \sum_{l=1}^{m}\sum_{k=1}^{K}\sum_{j_1,\dots,j_l=1}^{L} \alpha_{k0}\alpha_k \beta_{kj_1}\cdots\beta_{kj_l} E(X_{t-i}) \\
&\quad + \sum_{l=1}^{m}\sum_{k=1}^{K}\sum_{j_1,\dots,j_{l+1}=1}^{L} \alpha_k \alpha_{k_{j+1}} \beta_{kj_1}\cdots\beta_{kj_l} E(X_{t-j_1-\cdots-j_{l+1}} X_{t-i}) \\
&\quad + \sum_{k=1}^{K}\sum_{j_1,\dots,j_{m+1}=1}^{L} \alpha_k \beta_{kj_1}\cdots\beta_{kj_{m+1}} E(\lambda_{k(t-j_1-\cdots-j_{m+1})} X_{t-i})\,.
\end{aligned}$$

Using the same arguments as in the proof of Theorem 2.1, we can show that almost surely

$$\begin{aligned}
\gamma_{it} &= \sum_{k=1}^{K} \alpha_{k0}\alpha_k E(X_{t-i}) + \sum_{l=1}^{\infty}\sum_{k=1}^{K}\sum_{j_1,\dots,j_l=1}^{L} \alpha_{k0}\alpha_k \beta_{kj_1}\cdots\beta_{kj_l} E(X_{t-i}) \\
&\quad + \sum_{k=1}^{K}\sum_{j=1}^{L} \alpha_{kj}\alpha_k E(X_{t-j} X_{t-i}) \\
&\quad + \sum_{l=1}^{\infty}\sum_{k=1}^{K}\sum_{j_1,\dots,j_{l+1}=1}^{L} \alpha_k \alpha_{kj_{l+1}} \beta_{kj_1}\cdots\beta_{kj_l} E(X_{t-j_1-\cdots-j_{l+1}} X_{t-i}) \\
&= I + II + III + IV
\end{aligned}$$

with

$$\begin{aligned}
III &= \sum_{k=1}^{K}\sum_{j=1}^{L} \alpha_{kj}\alpha_k E(X_{t-j} X_{t-i}) \\
&= \sum_{k=1}^{K} \alpha_{ki}\alpha_k \gamma_{0,t-i} + \sum_{k=1}^{K}\sum_{j=1, i\neq j}^{L} \alpha_{kj}\alpha_k \gamma_{|j-i|,t} \\
&= \sum_{k=1}^{K} \alpha_{ki}\alpha_k \gamma_{0,t-i} \\
&\quad + \sum_{k=1}^{K} \alpha_k \Big(\sum_{|j-i|=1} \alpha_{ki}\gamma_{1,t} + \cdots + \sum_{|j-i|=i} \alpha_{kj}\gamma_{i,t} + \cdots + \sum_{|j-i|=L-1} \alpha_{kj}\gamma_{L-1,t} \Big) \\
&= \sum_{k=1}^{K} \alpha_k \delta_{i0k0}\gamma_{0,t-i} + \sum_{k=1}^{K}\sum_{u=1}^{L-1} \alpha_k \delta_{iuk0}\gamma_{u,t}
\end{aligned}$$

and

$$
\begin{aligned}
IV &= \sum_{l=1}^{\infty}\sum_{k=1}^{K}\sum_{j_1,\dots,j_{l+1}=1}^{L} \alpha_k \alpha_{kj_{l+1}} \beta_{kj_1}\cdots\beta_{kj_l}\gamma_{|i-j_1-\cdots-j_{l+1}|,t} \\
&= \sum_{l=1}^{\infty}\sum_{k=1}^{K}\sum_{j_1+\cdots+j_{l+1}=i}^{L} \alpha_k \alpha_{kj_{l+1}} \beta_{kj_1}\cdots\beta_{kj_l}\gamma_{0,t-i} \\
&\quad + \sum_{l=1}^{\infty}\sum_{k=1}^{K}\sum_{j_1+\cdots+j_{l+1}\neq i}^{L} \alpha_k \alpha_{kj_{l+1}} \beta_{kj_1}\cdots\beta_{kj_l}\gamma_{|i-j_1-\cdots-j_{l+1}|,t} \\
&= \sum_{l=1}^{\infty}\sum_{k=1}^{K} \alpha_k \delta_{i0kl}\gamma_{0,t-i} + \sum_{l=1}^{\infty}\sum_{k=1}^{K}\sum_{u=1}^{L-1} \alpha_k \delta_{iukl}\gamma_{u,t}
\end{aligned}
$$

where

$$
\delta_{iukl} = \sum_{|i-j_1-\cdots-j_{l+1}|=u} \alpha_{kj_{l+1}} \beta_{kj_1}\cdots\beta_{kj_l}.
$$

Then

$$
III + IV = \sum_{l=0}^{\infty}\sum_{k=1}^{K} \alpha_k \delta_{i0kl}\gamma_{0,t-i} + \sum_{l=0}^{\infty}\sum_{k=1}^{K}\sum_{u=1}^{L-1} \alpha_k \delta_{iukl}\gamma_{u,t}
$$

where the first term of this summation ($l = 0$) is III.

Moreover, using the same notation, we get

$$
\begin{aligned}
I + II &= \left(\sum_{k=1}^{K} \alpha_{k0}\alpha_k + \sum_{l=1}^{\infty}\sum_{k=1}^{K}\sum_{j_1,\dots,j_l=1}^{L} \alpha_{k0}\alpha_k \beta_{kj_1}\cdots\beta_{kj_l} \right)\mu \\
&= \left(\sum_{l=0}^{\infty}\sum_{k=1}^{K}\sum_{j_1,\dots,j_l=1}^{L} \alpha_{k0}\alpha_k \beta_{kj_1}\cdots\beta_{kj_l} \right)\mu =: K_1
\end{aligned}
$$

Finally, for $i = 1, ..., L$

$$
K_1 + \omega_{i0}\gamma_{0,t-i} + \sum_{u=1}^{L-1} \omega_{iu}\gamma_{u,t} = 0
$$

where

$$
\omega_{i0} = \sum_{l=0}^{\infty}\sum_{k=1}^{K} \alpha_k \delta_{i0kl}, \quad \omega_{iu} = \sum_{l=0}^{\infty}\sum_{k=1}^{K} \alpha_k \delta_{iukl}
$$

$$
\text{for} \quad u \neq i \quad \text{and} \quad \omega_{ii} = \sum_{l=0}^{\infty}\sum_{k=1}^{K} \alpha_k \delta_{iikl} - 1.
$$

Let $\Gamma = (\omega_{ij})_{i,j=1}^{L-1}$ and $\Gamma^{-1} = (b_{ij})_{i,j=1}^{L-1}$. The invertibility of the matrix Γ is checked in Appendix B.

Then

$$\Gamma(\gamma_{1,t},...,\gamma_{L-1,t})^T = -\big(K_1+\omega_{10}\gamma_{0,t-1},...,K_1+\omega_{(L-1)0}\gamma_{0,t-(L-1)}\big)$$

which is equivalent to

$$(\gamma_{1,t},...,\gamma_{L-1,t})^T = -\Gamma^{-1}\big(K_1+\omega_{10}\gamma_{0,t-1},...,K_1+\omega_{(L-1)0}\gamma_{0,t-(L-1)}\big).$$

We can show that

$$\gamma_{i,t} = -K_1\sum_{u=1}^{L-1} b_{iu} - \sum_{u=1}^{L-1} b_{iu}\omega_{u0}\gamma_{0,t-u}\,.$$

The second moment is given by:

$$\gamma_{0,t} \;=\; \mathbb{E}(X_t) + \sum_{k=1}^{K}\alpha_k\mathbb{E}(\lambda_{kt}^2)\,.$$

For $k=1,...,K$, we have

$$\begin{aligned}\lambda_{kt}^2 \;&=\; \Big(\alpha_{k0}+\sum_{i=1}^{L}\alpha_{ki}X_{t-i}+\sum_{j=1}^{L}\beta_{kj}\lambda_{k(t-j)}\Big)\lambda_{kt}\\ &=\; \alpha_{k0}\lambda_{kt}+\sum_{i=1}^{L}\alpha_{ki}X_{t-i}\lambda_{kt}+\sum_{j=1}^{L}\beta_{kj}\lambda_{k(t-j)}\lambda_{kt}\,.\end{aligned}$$

The hypothesis H_1 implies that the process $\{\lambda_{kt},\ t\in\mathbb{Z}\}$ is first-order stationary. Hence

$$\mathbb{E}(\lambda_{kt}) = \frac{\alpha_{k0}+\sum_{i=1}^{L}\alpha_{ki}\mu}{1-\sum_{j=1}^{L}\beta_{kj}} \quad \text{for } k=1,...,K.$$

We have

$$\begin{aligned}&\mathbb{E}\Big(\sum_{i=1}^{L}\alpha_{ki}X_{t-i}\lambda_{kt}\Big) \;=\\ &=\; \mathbb{E}\Big(C_{k0}\sum_{i=1}^{L}\alpha_{ki}X_{t-i}+\sum_{i=1}^{L}\alpha_{ki}X_{t-i}\sum_{l=0}^{\infty}\sum_{j_1,\dots,j_{l+1}=1}^{L}\alpha_{kj_{l+1}}\beta_{kj_1}\cdots\beta_{kj_l}X_{t-j_1-j_2-\cdots-j_{l+1}}\Big)\\ &=\; \mathbb{E}\Big(C_{k0}\sum_{i=1}^{L}\alpha_{ki}X_{t-i}+\sum_{l=0}^{\infty}\sum_{j_1,\dots,j_{l+2}=1}^{L}\alpha_{kj_{l+1}}\alpha_{kj_{l+2}}\beta_{kj_1}\cdots\beta_{kj_l}X_{t-j_1-j_2-\cdots-j_{l+1}}X_{t-j_{l+2}}\Big)\\ &=\; C_{k0}\mu\sum_{i=1}^{L}\alpha_{ki}+\sum_{i=1}^{L}\Delta_{k,i}^{(1)}\gamma_{0,t-i}+\sum_{v=1}^{L-1}\Lambda_{kv}^{(1)}\gamma_{v,t}\end{aligned}$$

where

$$\begin{aligned}\Delta_{k,i}^{(1)} \;&=\; \sum_{l=0}^{\infty}\sum_{\substack{j_{l+2}=i\\ j_{l+2}=j_1+\cdots+j_{l+1}}}^{L}\alpha_{kj_{l+1}}\alpha_{kj_{l+2}}\beta_{kj_1}\cdots\beta_{kj_l}\,,\\ \Lambda_{kv}^{(1)} \;&=\; \sum_{l=0}^{\infty}\sum_{|j_{l+2}-j_1-\cdots-j_{l+1}|=v}^{L}\alpha_{kj_{l+1}}\alpha_{kj_{l+2}}\beta_{kj_1}\cdots\beta_{kj_l}\,.\end{aligned}$$

Moreover

$$\begin{aligned}\sum_{j=1}^{L}\beta_{kj}\lambda_{k(t-j)} &= \sum_{j=1}^{L}\beta_{kj}\left\{C_{k0}+\sum_{l=0}^{\infty}\sum_{j_1,\dots,j_{l+1}=1}^{L}\alpha_{kj_{l+1}}\beta_{kj_1}\cdots\beta_{kj_l}X_{t-j-j_1-j_2-\cdots-j_{l+1}}\right\}\\ &= C_{k0}\sum_{j=1}^{L}\beta_{kj}+\sum_{l=0}^{\infty}\sum_{j_1,\dots,j_{l+2}=1}^{L}\alpha_{kj_{l+2}}\beta_{kj_1}\cdots\beta_{kj_{l+1}}X_{t-j_1-j_2-\cdots-j_{l+2}}\,.\end{aligned}$$

Hence

$$\begin{aligned}&\sum_{j=1}^{L}\beta_{kj}\lambda_{k(t-j)}\lambda_{kt} =\\ &= \left\{C_{k0}\sum_{j=1}^{L}\beta_{kj}+\sum_{l=0}^{\infty}\sum_{j_1,\dots,j_{l+2}=1}^{L}\alpha_{kj_{l+2}}\beta_{kj_1}\cdots\beta_{kj_{l+1}}X_{t-j_1-j_2-\cdots-j_{l+2}}\right\}\\ &\quad\times\left\{C_{k0}+\sum_{l=0}^{\infty}\sum_{j_1,\dots,j_{l+1}=1}^{L}\alpha_{kj_{l+1}}\beta_{kj_1}\cdots\beta_{kj_l}X_{t-j_1-j_2-\cdots-j_{l+1}}\right\}\\ &= C_{k0}^2\sum_{j=1}^{L}\beta_{kj}+C_{k0}\sum_{j=1}^{L}\beta_{kj}\sum_{l=0}^{\infty}\sum_{j_1,\dots,j_{l+1}=1}^{L}\alpha_{kj_{l+1}}\beta_{kj_1}\cdots\beta_{kj_l}X_{t-j_1-j_2-\cdots-j_{l+1}}\\ &+ C_{k0}\sum_{l=0}^{\infty}\sum_{j_1,\dots,j_{l+2}=1}^{L}\alpha_{kj_{l+2}}\beta_{kj_1}\cdots\beta_{kj_{l+1}}X_{t-j_1-j_2-\cdots-j_{l+2}}\\ &+\sum_{\substack{l=0\\ l'=0}}^{\infty}\sum_{\substack{j_1,\dots,j_{l+2}=1\\ j'_1,\dots,j'_{l'+1}=1}}^{L}\alpha_{kj_{l+2}}\beta_{kj_1}\cdots\beta_{kj_{l+1}}\alpha_{kj'_{l'+1}}\beta_{kj'_1}\cdots\beta_{kj'_{l'}}X_{t-j_1-j_2-\cdots-j_{l+2}}X_{t-j'_1-j'_2-\cdots-j'_{l'+1}}.\end{aligned}$$

The term $\mathbb{E}\Big(\sum_{j=1}^{L}\beta_{kj}\lambda_{k(t-j)}\lambda_{kt}\Big)$ is given by

$$\begin{aligned}\mathbb{E}\Big(\sum_{j=1}^{L}\beta_{kj}\lambda_{k(t-j)}\lambda_{kt}\Big) &= C_{k0}^2\sum_{j=1}^{L}\beta_{kj}+2C_{k0}\mu\sum_{l=0}^{\infty}\sum_{j_1,\dots,j_{l+2}=1}^{L}\alpha_{kj_{l+2}}\beta_{kj_1}\cdots\beta_{kj_{l+1}}\\ &\quad+\sum_{i=1}^{L}\Delta_{k,i}^{(2)}\gamma_{0,t-i}+\sum_{v=1}^{L-1}\Lambda_{kv}^{(2)}\gamma_{v,t}\end{aligned}$$

where

$$\begin{aligned}\Delta_{k,i}^{(2)} &= \sum_{\substack{l=0\\ l'=0}}^{\infty}\sum_{\substack{j_1+\cdots+j_{l+2}=i\\ j_1+\cdots+j_{l+2}=j'_1+\cdots+j'_{l'+1}}}^{L}\alpha_{kj_{l+2}}\beta_{kj_1}\cdots\beta_{kj_{l+1}}\alpha_{kj'_{l'+1}}\beta_{kj'_1}\cdots\beta_{kj'_{l'}}\\ \Lambda_{kv}^{(2)} &= \sum_{\substack{l=0\\ l'=0}}^{\infty}\sum_{\left|j_1+\cdots+j_{l+2}-j'_1-\cdots-j'_{l'+1}\right|=v}^{L}\alpha_{kj_{l+2}}\beta_{kj_1}\cdots\beta_{kj_{l+1}}\alpha_{kj'_{l'+1}}\beta_{kj'_1}\cdots\beta_{kj'_{l'}}.\end{aligned}$$

Let $\Delta_{k,i} = \Delta_{k,i}^{(1)} + \Delta_{k,i}^{(2)}$ and $\Lambda_{kv} = \Lambda_{kv}^{(1)} + \Lambda_{kv}^{(2)}$.

For $k = 1, ..., K$, the expectation of λ_{kt}^2 is given by

$$\mathbb{E}(\lambda_{k,t}^2) = C_k + \sum_{i=1}^{L} \Delta_{k,i}\gamma_{0,t-i} + \sum_{v=1}^{L-1} \Lambda_{kv}\gamma_{v,t}$$

with

$$C_k = \alpha_{k0}\mathbb{E}(\lambda_{kt}) + C_{k0}\sum_{i=1}^{L}\alpha_{ki}\mu + C_{k0}^2\sum_{j=1}^{L}\beta_{kj} + 2C_{k0}\sum_{l=0}^{\infty}\sum_{j_1,...,j_{l+2}=1}^{L}\alpha_{kj_{l+2}}\beta_{kj_1}\cdots\beta_{kj_{l+1}}\mu\,.$$

Then

$$\begin{aligned}
\gamma_{0,t} &= \mu + \sum_{k=1}^{K}\alpha_k\left(C_k + \sum_{i=1}^{L}\Delta_{k,i}\gamma_{0,t-i} + \sum_{v=1}^{L-1}\Lambda_{kv}\gamma_{v,t}\right)\\
&= \mu + \sum_{k=1}^{K}\alpha_k\Bigg[C_k + \sum_{u=1}^{L}\Delta_{k,u}\gamma_{0,t-u}\\
&\quad + \sum_{v=1}^{L-1}\Lambda_{kv}\left(-K_1\sum_{u=1}^{L-1}b_{vu} - \sum_{u=1}^{L-1}b_{vu}\omega_{u0}\gamma_{0,t-u}\right)\Bigg]\\
&= c_0 + \sum_{k=1}^{K}\alpha_k\left[\sum_{u=1}^{L}\Delta_{k,u}\gamma_{0,t-u} - \sum_{u=1}^{L-1}\left(\sum_{v=1}^{L-1}\Lambda_{kv}b_{vu}\omega_{u0}\right)\gamma_{0,t-u}\right]
\end{aligned}$$

where

$$c_0 = \mu + \sum_{k=1}^{K}\alpha_k C_k - K_1\sum_{k=1}^{K}\alpha_k\sum_{v=1}^{L-1}\Lambda_{kv}\sum_{u=1}^{L-1}b_{vu}.$$

Hence

$$\gamma_{0,t} = c_0 + \sum_{k=1}^{K}\alpha_k\left[\sum_{u=1}^{L-1}\left(\Delta_{k,u} - \sum_{v=1}^{L-1}\Lambda_{kv}b_{vu}\omega_{u0}\right)\gamma_{0,t-u} + \Delta_{k,L}\gamma_{0,t-L}\right]. \tag{5.1}$$

Let

$$c_u = \sum_{k=1}^{K}\alpha_k\left(\Delta_{k,u} - \sum_{v=1}^{L-1}\Lambda_{kv}b_{vu}\omega_{u0}\right),\quad u = 1, ..., L-1 \quad\text{and}\quad c_L = \sum_{k=1}^{K}\alpha_k\Delta_{k,L}\,.$$

Then the equation (5.1) is equivalent to:

$$\gamma_{0,t} = c_0 + \sum_{u=1}^{L}c_u\gamma_{0,t-u}\,. \tag{5.2}$$

The necessary and sufficient condition for a non-homogeneous difference equation (5.2) to have a stable solution, which is finite and independent of t, is that all roots of the equation: $1 - c_1Z^{-1} - c_2Z^{-2} - \cdots - c_LZ^{-L} = 0$ lie inside the unit circle.

APPENDIX B — Invertibility of Γ and positivity of c_0

In the following lines, we establish the invertibility of Γ and check the positivity of c_0. The same ideas were already used in the paper by Gonçalves *et al.* (2013).

Invertibility of Γ:

We show that the matrix $\Gamma = (\omega_{ij})_{i,j=1}^{L-1}$ is strictly diagonally dominant by rows. For $i = 1, ..., L-1$,

$$|\omega_{ii}| - \sum_{\substack{u=1\\u\neq i}}^{L-1} |\omega_{iu}| = 1 - \sum_{l=0}^{\infty}\sum_{k=1}^{K} \alpha_k \delta_{iikl} - \sum_{\substack{u=1\\u\neq i}}^{L-1}\sum_{l=0}^{\infty}\sum_{k=1}^{K} \alpha_k \delta_{iukl}$$
$$= 1 - \sum_{u=1}^{L-1}\sum_{l=0}^{\infty}\sum_{k=1}^{K} \alpha_k \delta_{iukl}\,.$$

We have

$$\sum_{u=1}^{L-1}\sum_{l=0}^{\infty}\sum_{k=1}^{K} \alpha_k \delta_{iukl} = \sum_{u=1}^{L-1}\sum_{l=0}^{\infty}\sum_{k=1}^{K} \alpha_k \sum_{|i-j_1-\cdots-j_{l+1}|=u} \alpha_{kj_{l+1}} \beta_{kj_1} \cdots \beta_{kj_l}$$
$$\leq \sum_{j_1,\ldots,j_{l+1}=1}^{L} \sum_{l=0}^{\infty}\sum_{k=1}^{K} \alpha_k \alpha_{kj_{l+1}} \beta_{kj_1} \cdots \beta_{kj_l}\,.$$

Based on the necessary condition for first-order stationarity in equation (2.5), we have

$$\sum_{j_1,\ldots,j_{l+1}=1}^{L} \sum_{l=0}^{\infty}\sum_{k=1}^{K} \alpha_k \alpha_{kj_{l+1}} \beta_{kj_1} \cdots \beta_{kj_l} < 1\,.$$

Hence, $|\omega_{ii}| - \sum_{\substack{u=1\\u\neq i}}^{L-1} |\omega_{iu}| > 0$. Then Γ is strictly diagonally dominant by rows. Hence, the matrix Γ is invertible by using the Levy–Desplanques Theorem (see Horn and Jonhson (2013), pp. 352, 392).

Positivity of c_0:

$$c_0 = \mu + \sum_{k=1}^{K} \alpha_k C_k - K_1 \sum_{k=1}^{K} \alpha_k \sum_{v=1}^{L-1} \Lambda_{kv} \sum_{u=1}^{L-1} b_{vu}\,.$$

To prove the positivity of c_0, it suffices to show that $b_{vu} \leq 0, v = 1, ..., L-1, u = 1, ..., L-1$. Indeed, it is easily seen that $-\Gamma$ is strictly diagonally dominant by rows. In addition, $-\omega_{ij} < 0$ for $i \neq j$ and $-\omega_{ii} > 0$ for $i = 1, ..., L-1$. Then $-\Gamma$ is a nonsingular M-matrix (see Quarteroni *et al.* (2000), p. 30, Property 1.20). This implies that $-\Gamma$ is inverse-positive that is $(-\Gamma)^{-1} \geq 0$. Hence, $\Gamma^{-1} \leq 0$, therefore $b_{vu} \leq 0$ for $v = 1, ..., L-1,\ u = 1, ..., L-1$.

ACKNOWLEDGMENTS

We wish to thank the referee as well as the associate editor for helpful comments and suggestions, which have permitted to improve the presentation of this paper.

REFERENCES

[1] Dempster, A.P.; Laird, N.M. and Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm, *J. Roy. Statist. Soc.*, **Series B 39**, 1–38.

[2] Der, G. and Everitt, B.S. (2002). *A Handbook of Statistical Analysis using SAS, second ed.*, Chapman and Hall/CRC, London.

[3] Ferland, R.; Latour, A. and Oraichi, D. (2006). Integer-valued GARCH process, *J. Time Ser. Anal.*, **27**, 923–942.

[4] Goldberg, S. (1958). *Introduction to Difference Equation*, Wiley, New York.

[5] Gonçalves, E.; Mendes-Lopes, N. and Silva, P. (2013). Infinitely divisible distributions in integer valued GARCH models, *Preprint Number 13-38*, Universidade de Coimbra.

[6] Gradshteyn, I.S. and Ryzhik, I.M. (2007). *Table of Integrals, Series, and Products*, seventh ed., Academic Press, San Diego.

[7] Horn, R.A. and Jonhson, C.R. (2013). *Matrix Analysis*, Cambridge University Press, New York, 2nd ed.

[8] Karlis, D. and Xekalaki, E. (2003). Choosing initial values for the EM algorithm for finite mixtures, *Comput. Statist. and Data Anal.*, **41**, 577–590.

[9] Martin, V.L. (1992). Threshold time series models as multimodal distribution jump processes, *J. Time Ser. Anal.*, **13**, 79–94.

[10] Melnykov, V. and Melnykov, I. (2012). Initializing the EM algorithm in Gaussian mixture models with an unknown number of components, *Comput. Statist. and Data Anal.*, **56**, 1381–1395.

[11] Muller, D.W. and Sawitzki, G. (1991). Excess Mass Estimates and Tests for Multimodality, *J. Amer. Statist. Assoc.*, **86**, 738–746.

[12] Naik, P.A.; Shi, P. and Tsai, C.L. (2007). Extending the Akaike information criterion to mixture regression models, *J. Amer. Statist. Assoc.*, **102**, 244–254.

[13] Nityasuddhia, D. and Böhning, D. (2003). Asymptotic properties of the EM algorithm estimate for normal mixture models with component specific variances, *Comput. Statist. and Data Anal.*, **41**, 591–601.

[14] Quarteroni, A.; Sacco, R. and Saleri, F. (2000). *Numerical Mathematics*, Texts in Applied Mathematics 37, Springer-Verlag, New York.

[15] Saikkonen, P. (2007). Stability of mixtures of vector autoregressions with autoregressive conditional heteroskedasticity, *Stat. Sinica*, **17**, 221–239.

[16] Titterington, D.M.; Smith, A.F.M. and Markov, U.E. (1985). Statistical Analysis of Finite Mixture Distributions, Wiley, New York.

[17] Wong, C.S. and Li, W.K. (2000). On a mixture autoregressive model, *J. Roy. Statist. Soc.*, **Series B 62**, 95–115.

[18] Zeger, S.L. (1988). A regression model for time series of counts, *Biometrika*, **75**, 621–629.

[19] Zhu, F. (2011). A negative binomial integer-valued GARCH model, *J. Time Ser. Anal.*, **32**, 54–67.

[20] Zhu, F. (2012). Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued GARCH models, *J. Math. Anal. Appl.*, **389**, 58–71.

[21] Zhu, F.; Li, Q. and Wang, D. (2010). A mixture integer-valued ARCH model, *J. Statist. Plann. Inference*, **140**, 2025–2036.